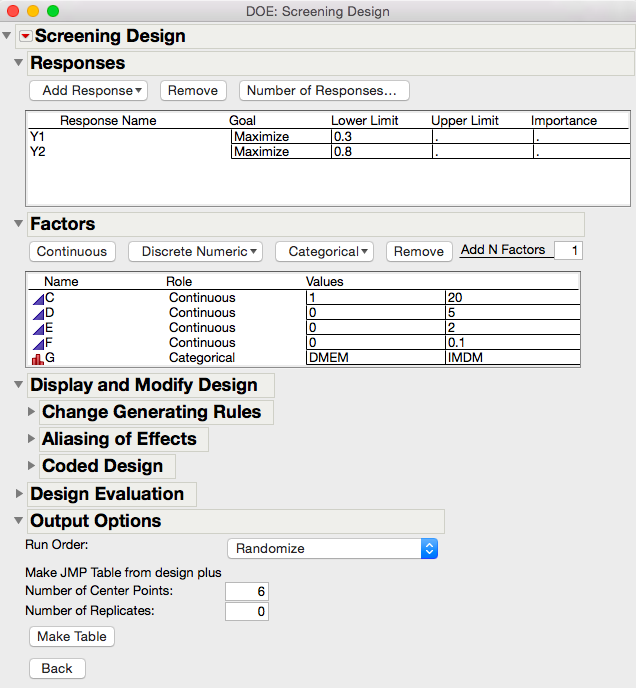
**Experimental Strategy:**

**Request 1: Fractional Factorial Screening Design**

For the first data request, we used a fractional factorial to screen the effects of factors A, B, C, D, E, F, G on Y1 and Y2. For each batch request, we were limited to 70 runs. For a full factorial design of 7 factors, we would need 2^7 = 128 runs, which was not possible therefore we did a one half fraction of the 7 factor factorial design, so we could maximize the number of runs in the screen. We requested 2^7-1 = 64 run factorial with six centre points to determine the curvature of the regression model.

What are the aliases?



**Request 2: Central Composite Inscribed Design**

After analysis of the screen, we narrowed found that factors A, B, C, and D were significant, as per ANOVA described below, and we decided to do a rotatable central composite inscribed design with six centre points.

ccd: alpha = (2^k)^0.25 = 2, 6 centre points, CCD rotatable, inscribed

Since we expect to have a response surface model, a rotatable CCD is a desirable property for the quadratic model design as the variance of the predicted response at any point x depends only on the distance from x to the center point. Also, we chose CCI (inscribed) because the axial points cannot exceed the +1, -1 limits of the design (which are the maximum and minimum levels of A, B, C, D, E, F.

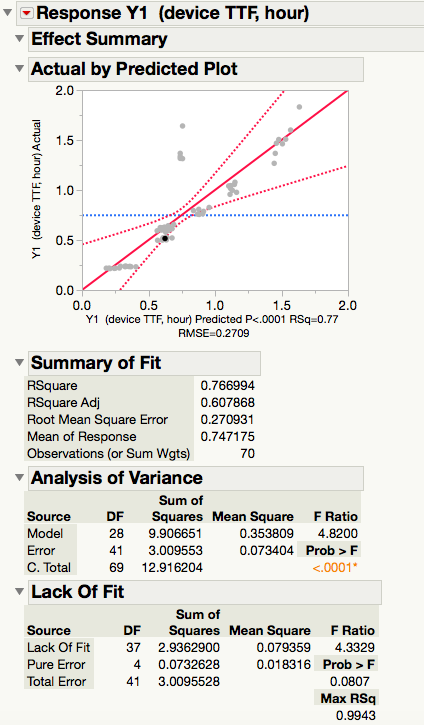
Alpha = [number of factorial runs]^1/4 = [2^k]^0.25 for a rotatable design

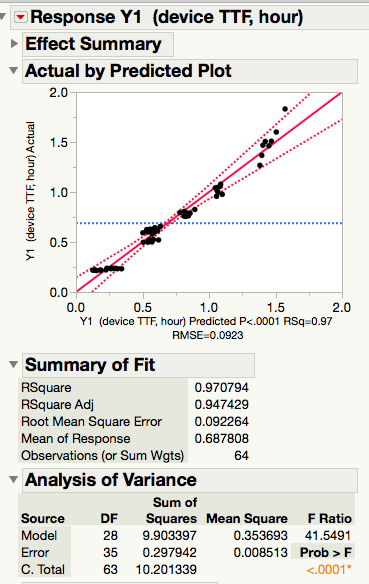
**Identify the Significant Factor Effects and Interactions:**

Data Analysis 1: (Screen Test)

|  |  |  |
| --- | --- | --- |
| Factor or Interaction | Effect on Y1 | Effect on Y2 |
| A | - | - |
| B | + | + |
| C | - | + |
| D | none | + |
| AB | - | None |
| AC | + | None |
| BC | - | none |
| AD | None | - |
| CF | none | + |

Y1

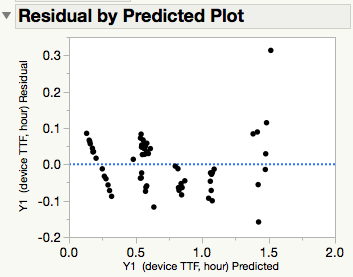
  
Shows Y1 fit model. The lack of fit is not significant, however it does exist and the “Actual by Predicted” plot has points that deviate from the line, and outside the dotted red lines, and Rsquare was 0.767 which could be improved. We hypothesized this was due to the centre points, as they help model curvature which would deviate from a linear model, so they were removed from the data. The objective of this part of the experiment is to narrow down the significant factors, therefore curvature and the actual predicted model which is not linear can be modeled later.



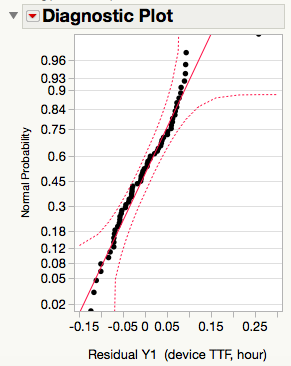
Here, the Actual by Predicted plot has an Rsquare of 0.97 and the ANOVA shows the model is significant therefore a good linear regression model to fit this fractional factorial design data without centre points. There is no longer a lack of fit.

Therefore the significant factors for Y1 are A, B, C, and interactions are AB, AC and BC.

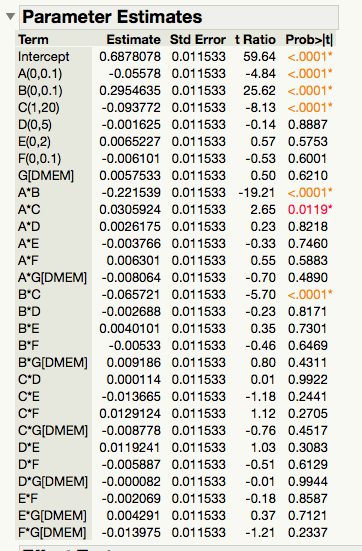
Residuals for Y1 no centre points linear regression model:



There does not seem to be a pattern, there may be a slight funnel here, but I think that it’s just due to the outlier at the top right.



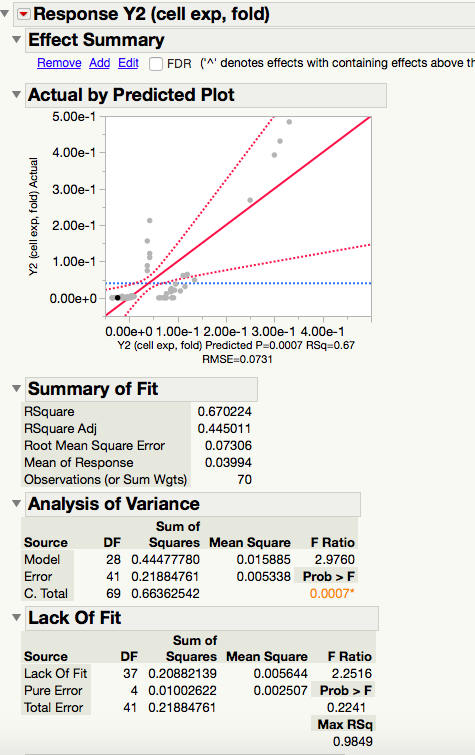
From this plot, it looks like the residuals are normally distributed (is there a way we can show rsquared?)



Thus the assumptions for ANOVA are satisfied for the no centre points Y1 linear regression model:

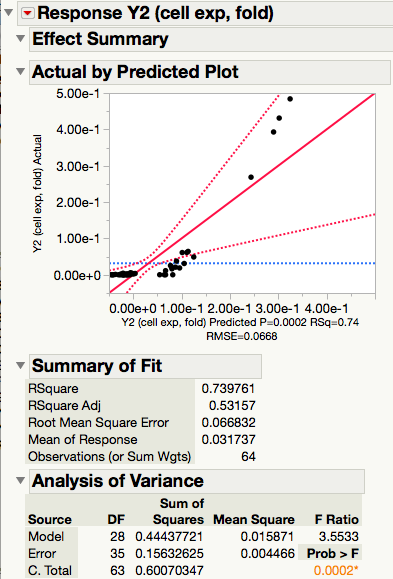
* Residuals are normally distributed
* Independent observations were requested
* How do we determine equal variance?

Y2

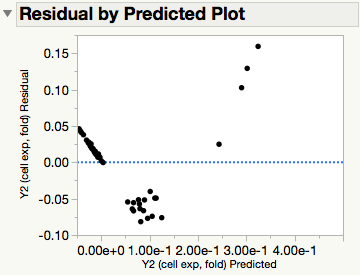


This was the linear regression model with centre points for Y2. The Actual by Predicted Plot has a low Rsquare value of 0.67 and a lack of fit, which is insignificant but the plot does not look good because the points are not within the line or the dashed lines.

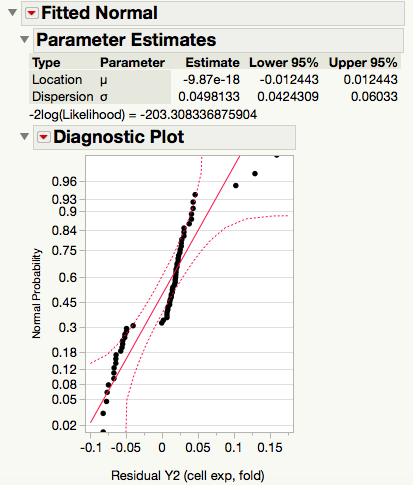
Therefore, for the same reason as above for Y1, the centre points were removed from the data.



Here, the Rsquare is still low at 0.74. There is no lack of fit but the Actual by Predicted plot does not look like a good model plot.



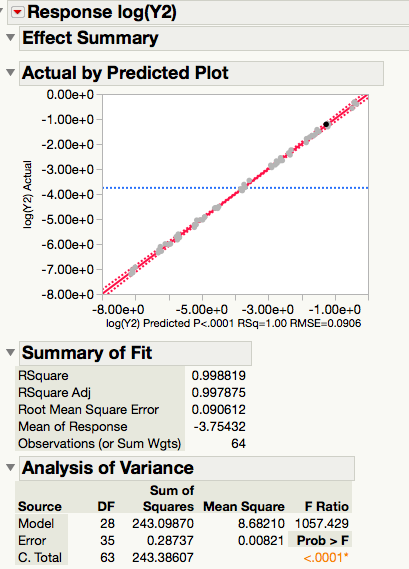
Here, the residuals seem to follow some sort of pattern and are not randomn.



By the normal probability distribution plot, it does not look like the residuals are normally distributed.

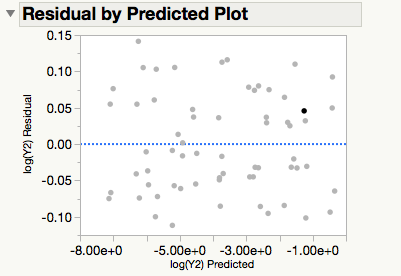
Since the residuals do not look good, we transformed Y2 to a log10(Y2)

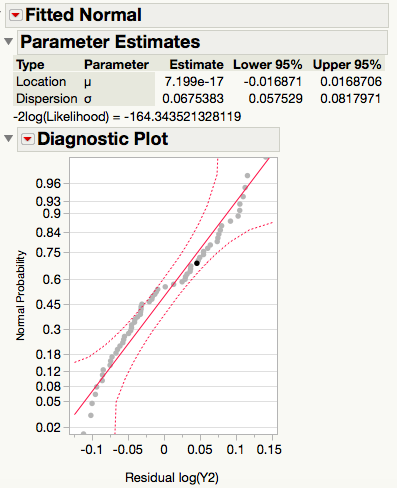
Now, when the linear regression model was fitted to log10(Y2):

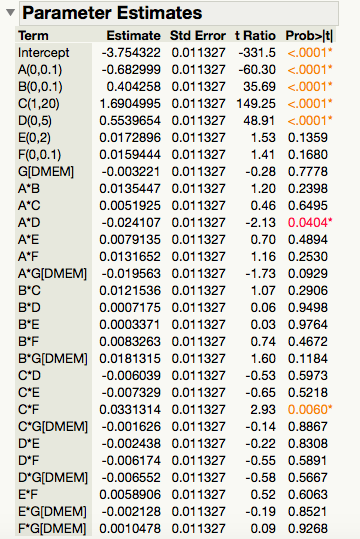


The Rsquare is 0.9998 for the “Actual by Predicted Plot”, the ANOVA for the model is significant.

Residual Analysis:







Therefore the significant factors for Y2 are: A, B, C, D and interactions AD, CF

Data Analysis 2: (Response Surface Model)

|  |  |  |
| --- | --- | --- |
| Factor | Effect on Y1 | Effect on Y2 |
| A |  |  |
| B |  |  |
| C |  |  |
| D |  |  |

Analyze data Batch 1 – Fit Model

exclude centre points -> fit model for 3 interactions

Use diagnostic tests – anova and t test

Ensure assumptions are satisfied

Find proper method to analyze data satisfying the assumptions and do all the proper tests

Then request data

Evaluate design – colour map on correlation

For ccd determination: what is the best alpha? Use the colour map on correlation

* Can use for example alpha = 1.8 - > set -1 = -1.8 and 1 = 1.8 and recode the values

ANOVA Assumptions = normally distributed residuals, observations are independent, variance is the same for all groups

Linear regression assumptions include:

* the means of response variable are accurately modeled by a linear function of the factors.
* The random error term is assumed to be normally distributed with a mean of zero and constant variance
* Errors are independent

By analyzing the residuals to determine model adequacy, the following assumptions can be checked:

* Errors are approximately normally distributed with constant variance
* If data transformation or additional terms in the model would be useful

This is what the Y2 residuals look like:

We did a log transform because before the residuals did not look good. Therefore, when we transform Y2 to log10Y2, and here is the corresponding residual analysis:

Is the F ratio affected when transformed?

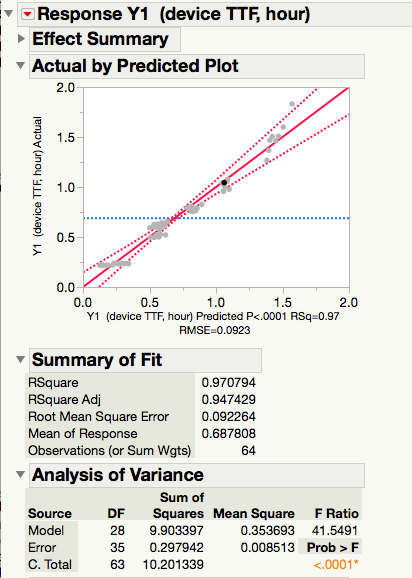
Y2 significant factors with linear regression =

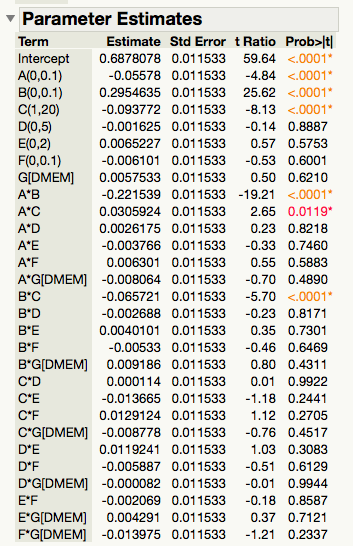
Y2 significant factors with logy2 transformed linear regression =

In the transformation: What is alpha? What is the relationship between mu and sigma? To determine the transformation see ch 15 slide 17 for the table

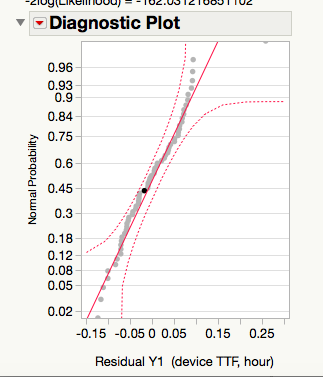
Y1 Fit Model

Removed centre points

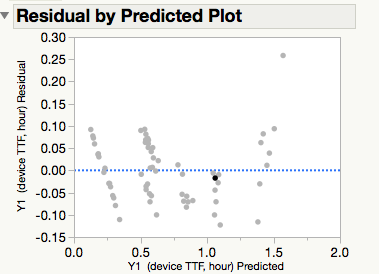




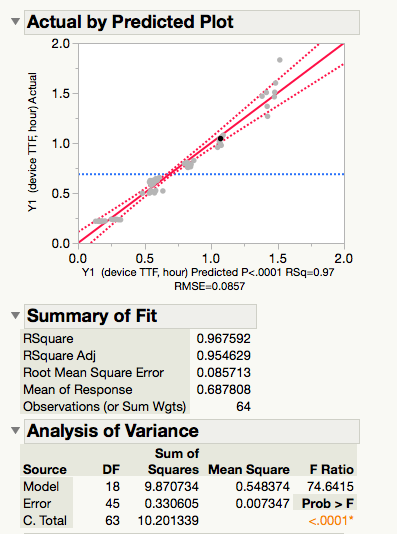
Therefore, A, B, C are significant factors and AB, AC, BC are the significant interactions

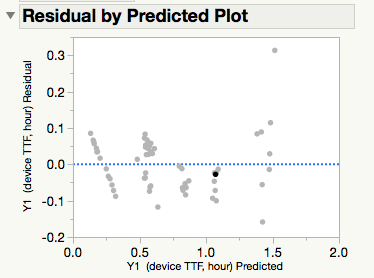


therefore residuals are normally distributed – satisfying ANOVA assumption

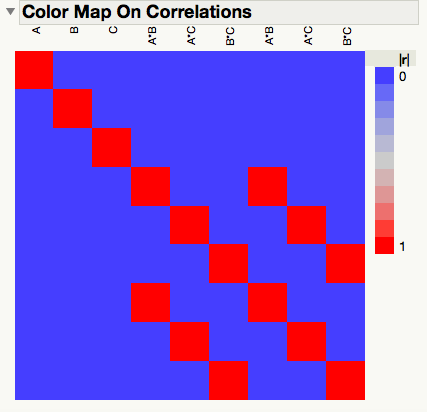


Y1 analysis with significant factors (ABC):





colour map on correlation for significant factors:



no correlation between factors

Finding the response suface – will generate a model that can be used to interpolate the best factor levels to achieve the best trade off for the factor goals

Factor goals: Y2

Y1:

Factor constraints:

After requesting data\_2,

Log(y2)

Recip(y1)

Show that when only the significant factors and interactions are fitted then the model is adequate (no lack of fit)

Fit model both log(y2) and recip(y1) together and flipped log(y1) so that when you maximize desirability it is actually minimizing it

That generates optimal factor levels for maximizing log(y2) and minimizing recip(y1)

Put that into the contour profiler and the cross hair shows where those factor levels are on the combined contour profiler

Then put high = 3.33 for recip(y1)

Low = -0.096 for log(y2) and that generates shaded regions

If the cross hair is in the white region then the factor levels are in the intended high/low level region and so it is showing the optimal levels for the factors to optimize both responses.